

Operator Localization of Virtual Objects

Stephen R. Ellis
NASA Ames Research Center
Moffett Field, California, 94035 USA

Brian M. Menges
San José State University Foundation
San José, California, 95192 USA

Abstract

Errors in the localization of nearby virtual objects presented via see-through, helmet mounted displays are examined as a function of viewing conditions and scene content. Monocular, biocular or stereoscopic presentation of the virtual objects, accommodation (required focus), subjects' age, and the position of physical surfaces are examined. Nearby physical surfaces are found to introduce localization errors that differ depending upon the other experimental factors. The apparent physical size and transparency of the virtual objects and physical surfaces respectively are also influenced by their relative position when superimposed. Design implications are discussed.

1. Introduction

The following four experiments explore such phenomenon by examining subjects' ability to adjust the distance of a physical pointer to match that of a nearby virtual object. The object is generated by a high performance, computer graphics system and presented by a head-mounted, see-through display. This localization task was selected since it is close to the visual-manual manipulation expected of users of virtual objects in numerous possible applications from surgery to mechanical assembly on a production line. In these applications virtual objects are proposed to become 3D guides that instruct and aid manual interaction with the world by dynamically helping the user position his hand. Thus, these virtual objects must be accurately and reliably localized by the user. For this reason, and because its precision and accuracy can be easily and reliably measured, the following experiments study the localization of virtual objects with a real pointer. Preliminary testing showed that our subjects could set a mechanically displaced, physical pointer to match the distance of physical targets with several millimeter accuracy and that this accuracy corresponded to their ability to match target distances with their fingers. Most previous work has examined system configuration to optimize display performance (e.g. [1] [2]) but only Rolland's [3] laboratory has systematically investigated user localization performance.

The initial experiment below examines the effects of three different viewing conditions on the subjects' accuracy of placement of the physical pointer under a virtual object. The conditions are: 1) monocular, 2) biocular or 3) stereoscopic viewing. These represent a range of cost and image fidelity, i.e. the completeness with which an

objects physical characteristics are presented. The monocular condition presents a virtual image of an object much as monocular helmet mounted sights do and represents a minimal hardware/software rendering cost for such virtual image displays. But the monocular virtual images it presents are subject to visual suppression due to binocular rivalry. The biocular condition, which presents two identical virtual images to the subjects' left and right eyes, can avoid the rivalry problem. The images it presents are projected from a cyclopean position between the viewers eyes but are shifted to allow unstressed fusion at a selected distance (See Figure 1). This condition halves the rendering cost with respect to stereo displays, but doubles the head-mounted display hardware requirements with respect to monocular displays. It presents a pattern of disparity approximating that of a transparent flat display surface and represents an intermediate cost system with potentially more stable image brightness. The third condition, conventional stereo display with parallel viewing vectors, presents the highest spatial fidelity but doubles both the rendering and hardware display cost compared to the monocular condition.

Because the monocular and biocular viewing conditions degrade the fidelity with which distance is presented, we expect that the stereoscopic display should support the most accurate localization. The biocular display may, however, provide a competitive alternative for virtual objects with relatively little internal depth by avoiding the potential problem of binocular rivalry. The first experiment provides a descriptive study of relative localization accuracy of virtual objects presented via the alternative viewing conditions. Since observer age and accommodative demand (required focus) could be expected to interact and influence localization accuracy, these display characteristics are also examined in the first study to develop designer guidelines for the adjustment of focus and selection of personnel to use virtual object displays. Subsequent studies consider the effect of the introduction of a nearby physical surfaces on the localization of the virtual object, identify a phenomenon that introduces errors into such localization, and explore two alternative explanations for the phenomenon.

2. Experiment 1

2.1 Methods

All experiments reported in this paper used a custom, head-mounted, see-through display, called an electronic

haploscope by the authors, capable of presenting a 20-30° diameter circular monocular field to each eye with variable monocular overlap. In the following experiments the system was used at 100% overlap and 20° field of view. The display system used two vertically mounted Citizen 1.5' 1000 line miniature Cathode Ray Tube's (CRT) in National Television Standards Committee (NTSC) mode which were driven by a Silicon Graphics (SGI) computer (4D/210GTXB) through custom video conditioning circuits. For the simple 3D imagery used in the following experiments, the computer could maintain a 15 Hz graphics update rate. The CRT images were infinity collimated by standard glass telescope eyepieces (Erf1 32 mm and Ploessl 42 mm) mounted directly under the CRTs. After the signal transformation from the RGB to NTSC, individual pixels which corresponded in the current configuration to 3 arcmin horizontal resolution measured from subjects' eyes were easily discriminated. The collimated light could be modified by lenses and rotating prisms from a standard optometric trial lens set that allowed precise positioning with at least 5 arcmin resolution of the separate left and right images and allowed variation of the accommodative demand for each eye. The images were relayed to the subjects eyes by custom, partially silvered (15%) polycarbonate mirrors mounted at 45° directly in front of each eye. The left and right viewing channels could be mechanically adjusted between 55 mm and 71 mm separations for different subject's interpupillary distances. The video signal conditioning also allowed lateral adjustment of the video frame. Consequently, the display system can precisely position the center of each graphics viewport directly in front of the eyes of all subjects for bore-sight alignment.

The entire display system, built around a snug fitting, rigid headband, is intended to be worn by a freely moving subject and weights between 0.77 and 1.1 kg. depending upon configuration. In the lightest configuration the moments of inertia have been measured when mounted on an erect head to be: 0.0782 kg-m² vertical axis, 0.0644 kg-m² longitudinal axis, 0.0391 kg-m² lateral axis. In all of the experiments described below, the band was fitted to each subjects head and then supported by a special pivoted mount at the end of a 1.8 m table. This mount restricted horizontal movement but allowed some pitch movement. Subjects sat at this end during the course of each experiment. The mount and chair were adjusted so that the virtual objects could be presented at eye level. Lateral head movement was restricted during all the following experiments but a residual pitch of $\pm 10^\circ$ was allowed for subject comfort. In practice the subjects were monitored by the experimenters to keep their heads approximately level at an individually selected orientation during the course of the experiments.

A monocular, biocular or stereoscopic virtual image of an upside down, axially rotating (~3 rpm) pyramid was presented at a distance of 58 cm away from the subjects' eyes by a head-mounted see-through display. It was seen

against a grey, cloth covered wall 2.2 m from the subjects. Preliminary experiments examining varying the rotation rate of the pyramid for each trial showed that such variation had no effect on the localization of the virtual image. The reference distance of 58 cm was chosen for the experiment because it corresponded to a possible working distance for several industrial applications of interest to the authors. All displays were operated under moderate indoor artificial illumination (approximately 50 lux). The virtual image was presented with either 2 diopters accommodative relief or at optical infinity. The stereo display was, however, calibrated (see below) over a range of 30 to 110 cm. The monocular display was simply the stereo channel that corresponded to the subjects' dominant eye. The biocular display was produced by positioning the graphics eye point midway between the subjects eyes. The left and right images were identical copies of this view but were shifted laterally so that when the subjects eyes converged to the centers of each view-port they would have 0 disparity relative to the reference convergence point 58 cm. The plane of 0 disparity was thus set so that the subjects could easily fuse the images when converged at 58 cm. This technique was used in general for all biocular stimuli at different depths which were experimentally interjected as described below. Though no keystone correction was applied for the distortion caused by the image shift, the disparity pattern produced in this biocular image closely approximates that of a flat image of the target as if it were drawn on a transparent projection surface at the simulated convergence distance.

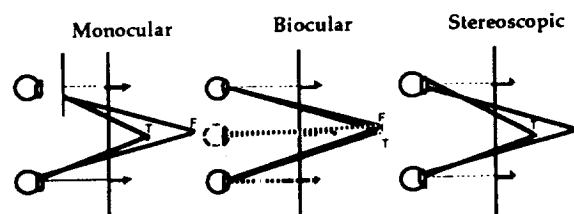


Figure 1. Viewing geometry for each viewing condition.

The wire-frame pyramid had a nominal 10 cm base and 5 cm height. The width of the wire frame lines were about 9 arcmin. The depicted size of the presented virtual object was randomly scaled from 70 to 130% of its nominal size for each trial to interfere with subjects' possible use of angular size as a depth cue. The lines of the wire frame and all other computer generated lines had a luminance of about 65 cd/m² and were seen against the gray cloth background of 2.9 cd/m² with visible vertical seams. While the presented luminance did not approach that used in aircraft heads up displays which must be visible against a 30,000 cd/m² background, it was just adequate for indoor work against most colored surfaces and is at least 3 times brighter than the luminance available in off-the-shelf, see-through head-mounted displays such as the IGlasses™ formerly made by VIO.

Four out of every 30 judgments were based on unanalyzed random variations in the depicted depth of the pyramid. This variation was introduced to help insure that the subjects did not notice that the same depicted depth was repeated. However, the major factor masking the repetition of the depicted depth were the perceptual effects causing changes in the apparent depth with different viewing conditions. Since the viewing conditions could be unobtrusively intermixed and no feedback was given to the subjects, there was no way for them to tell that the depicted distance was not in fact changing. In fact, no naive subject in any of the following experiments reported noticing the repetition of the target depth.

The subjects' task was to use a method of adjustment to position the binocularly visible, physical cursor, a yellow-green light emitting diode (LED) (about 20 cd/m²) pointer, shaped like a pyramid (base 0.5 cm, height 1 cm) into vertical alignment with the apex of the inverted pyramidal virtual object. The physical cursor was moved on a rail by a chain and gear system and positioned about 2 degrees of visual angle below the virtual object. The distance to the pointer was automatically recorded through use of a shaft-encoder interfaced to the display computer. The adjustment was self paced, but subjects were encouraged to take between 15 and 30 seconds for each adjustment and were allowed to take breaks at one half to 1 hours intervals as needed. As part of the standard procedure for use of human subjects, all subjects were informed they could terminate the experiment at anytime if they experienced any undue discomfort and were asked at the end if they experienced any viewing difficulties seeing the virtual objects.

The haploscope display system was adjusted by monocular superimposition of reference virtual images on an 18 cm diameter circumscribed circle presented at a distance of 2.2 m. In addition to position adjustment this allowed adjustment of the field of view angle of the graphics system to match the total magnification of the system. This was done separately for each eye and for each subject before the experiment was started. Thus, we could account for any changes due to variation of accommodative demand and corrective lenses that might be worn by the subjects. The subjects' interpupillary distances (IPD) were measured with a binocular-type viewing device with digital readout (Varilux Model: Digital CRP). All displays and algorithms were adjusted to reflect the measured IPD values.

Preliminary test results for virtual targets placed between 33 and 108 cm showed that within the full range of adjustment used for the experiment subjects using a stereo display could align the cursor within ± 0.3 cm of the depicted virtual object target depth [4]. The distance responses were completely linear, unbiased and unskewed and were conducted in the same full room illumination as the experiment. Similar tests of the biocular viewing condition showed equally linear responses. Tests in the

monocular condition showed, expectedly, inconsistent behavior. In further examination of the localization technique pilot subjects were asked to use the pointer to match the depth of physical targets. These tests showed linear, unbiased, unskewed estimates with statistical ranges of ± 0.15 cm about depicted physical distances for targets used in the experiment.

Ten subjects, five young (15 - 29 yrs.) and five older (38 - 47 yrs.) participated in the experiment. Subjects in the older group could be presumed by population data to be at least early presbyopes [5]. All but one young and one older subject (i.e. the authors) were naive with respect to the purpose of the experiment. The others were either paid subjects recruited through the Ames Bionetics contractor or were laboratory personnel. All subjects were screened on the Bauch & Lomb Orthorater stereo tests for stereoacuity better than 1 arcmin. Subjects who normally wore prescription spectacles were allowed to wear them during the screening test and during the experiment. During pilot testing for the experiment, inadvertent errors of 0.1 - 0.2 cm in modeling of subjects' interpupillary distances in the graphics simulation produced easily noticeable artifacts. Precise stereo or biocular presentation of virtual objects evidently requires measurement and modeling of interpupillary distance with an accuracy on the order of ± 0.1 cm.

Viewing Conditions (monocular, biocular, and stereoscopic) were crossed with Accommodation (0 or 2 Diopters) and nested within Age groups. The experiment used a blocked design in which blocks of 5 replications of a given condition were presented for each of the three viewing conditions producing uninterrupted 15 judgment sequences. The sequence of viewing conditions were randomly assigned to each subject and thereafter systematically permuted after each set of 3 viewing conditions were presented. In general, it was possible to switch the viewing conditions solely through software. Thus, the subjects were generally unaware which viewing condition was presented and the perceptual variation in apparent depth caused by variation in view condition was readily interpreted by them as variation in depicted depth. The viewing conditions were blocked for a given accommodative demand which was switched by interrupting every 15 trials to change viewing lenses. The order of presentation of accommodative demand was permuted within subjects and balanced across subjects.

2.2 Results

Analysis of variance (ANOVA) showed that the viewing conditions had a major effect on the bias of the subjects distance judgments ($F = 15.580$, $df = 2, 16$, $p < .001$). The mean stereoscopic and biocular localizations were almost completely correct, but a judgment bias appeared as an overestimate when the stereo depth cues associated with the virtual object were removed by monocular viewing. This effect interacted with accommodative

demand and age, as indicated in Figure 2 ($F=7.76$, $df = 2$, 16 , $p < .004$). All other effects are related to this three way interaction and will not be discussed individually. No subjects reported any difficulties seeing the virtual objects during the experiment.

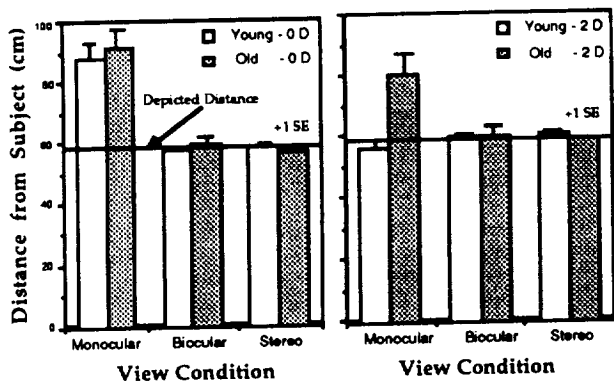


Figure 2: Effect of Age/Accommodative on localization.

2.3 Discussion

The results of the ANOVA plotted in Figure 2 show that when depth cues to the virtual object are degraded to monocular conditions, judgments of its distance drop back towards the distance of the background wall at 2.2 m from the subjects. A phenomenon that could explain the increased judged target distance in the monocular condition is the specific distance effect which causes visual targets of unknown physical size presented with weak or ambiguous depth cues to tend to appear in visually impoverished environments about 2-3 meters away [6]. This distance effect is also associated with tonic accommodation and vergence which relax to approximately 1-2 meters in the absence of distance cues [7]. Changes of convergence to a more distant resting position could cause the localization of the virtual object to recede from the viewer. But since these effects are generally seen when targets appear against featureless backgrounds they are not likely important for the present results.

Another phenomenon probably more relevant is the direct effect of change in ocular convergence on perceived depth [8] [9]. It is the kind of result to be expected if the subjects' actual convergence was driven or attracted by the wall which provided a visually sharp, textured cloth with vertical seams at 2.2 m. In the monocular condition, the only source of information that the virtual object is located any particular distance in front of the background is provided by accommodative demand. It is therefore not surprising that when 2 diopters of accommodative demand was provided, only the subjects young enough to respond to this cue were able to localize the virtual object approximately correctly. In these subjects the accommodative-convergence reflex allowed them to maintain convergence on the virtual object. Older subjects unable to respond to accommodative cues only have the disparity information

provided by the background cloth to control their vergence. Thus, they would still diverge to fuse the background and the monocularly presented virtual objects would still appear toward or on the background wall.

That the monocular virtual objects are not judged to be exactly on the wall, reflects the response bias within the experiment originating from constraints on movement of the physical cursor, unavoidable guesses the observers may have made about the approximate size of the object, and the possibility that the observer converged to other closer objects such as the pointer and its supports which, though darkened to be less conspicuous, were still visible against the background. In fact, if the observers were to hold their eyes completely still, the monocular viewing situation is very similar to that of viewing a monocular afterimage in a demonstration of Emmert's Law, the classic observation that an afterimage often appears at the distance of a physical surface against which it is projected [10]. The readily available correct disparity information in the binocular and stereo conditions, however, provide the missing cue that allowed all observers to correctly judge the distance to the virtual object.

Finally, since no difficulties seeing the virtual objects were reported during the experiment, we find no evidence that binocular rivalry interfered with subjects' ability to see the virtual objects in the monocular condition. The finding is consistent with all observations we have made of the monocular wire-frame virtual objects during preparation for the experiment. Apparently, the high contrast & motion of the objects we have examined easily overcomes any binocular rivalry that might be present.

3. Experiment 2

Experiment 1 examined the effect of different viewing conditions on the localization of virtual objects superimposed on a physical surface 2.2 m distant. But since new uses of virtual objects are likely to bring them closer to physical surfaces, Experiment 2 examines the effect of introduction of a much closer physical surface. In view of the interacting roles of accommodative and convergence in the discussion of Experiment 1, one could reasonably expect the introduction of a nearby real surface to cause the observers to localize monocularly viewed virtual objects at the same distance of introduced surface. If the accommodative demand for the virtual object is already matched to its displayed distance, one would expect that this improvement in the accuracy of localization would be larger for older observers than younger ones who already would have accommodative cues to the virtual object distance [11]. Accordingly, subjects of different age groups were used while correct accommodative demand to the virtual object was provided. In order to study the effects of introduction of a physical surface, subjects must first judge the distance of the virtual image by itself. This first judgment is identical to those made with the viewing conditions in Experiment 1 with 2 D accommodative de-

mand and provides a chance to replicate that part of the experiment.

3.1 Methods

The virtual image stimuli used in Experiment 2 were identical to those in Experiment 1 for the 2 diopter accommodation condition but a new physical stimulus was introduced. This physical surface was a slowly, irregularly rotating checkerboard (~2 rpm) made of xeroxed paper glued on foam-core and was mechanically introduced along the line of sight to the pyramid as illustrated in Figure 3. Motion of the checkerboard was introduced be-

cause preliminary testing showed that changes in localization that it produced were enhanced by motion. The checkerboard was a disk 29 cm in diameter with 5 cm black and white checks having either 1.3 cd/m² or 17.8 cd/m² luminance. It was positioned so that the virtual image of the pyramid could be seen against the lower rim of the disk in order to allow the subjects to adjust the physical cursor to the apparent distance of the virtual image in the presence of the disk. Care was taken to be sure the physical cursor was below the bottom of the virtual object and the edge of the disk as in Experiment 1. As before, subjects viewed the virtual objects with either monocular, biocular or stereoscopic view conditions.

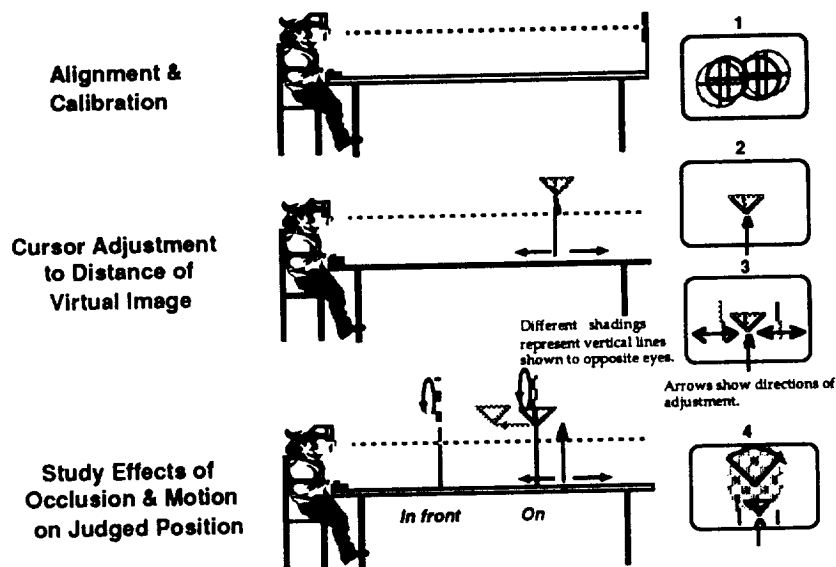


Figure 3: Experimental procedure illustration for Experiments 2, 3, and 4. Top: alignment, magnification, and interpupillary adjustment, Middle: Initial localization of virtual object depth, Bottom: Testing conditions representing, the "on" or "in front" placement of the rotating checkerboard and the second localization of the virtual object depth. The rightmost panels 3 and 4 represent use of nonius lines to detect relative convergence in Experiment 3.

Thirteen subjects, seven young (15 -29 yrs.) and six older (38 - 47 yrs.) participated in the experiment. All but one young and one older subject (the authors) were naive with respect to the purpose of the experiment. The others were either recruited through the Ames Bionetics contractor or were laboratory personnel. All subjects were screened for stereo vision as in Experiment 1.

The first part of the subjects' task was to mechanically place the yellow-green LED pointer under the nadir of the slowly rotating, wire-frame virtual pyramid, which varied randomly in size for each trial as in Experiment 1. The second part of the task involved an adjustment of the pointer to match the pyramid's distance after the slowly, irregularly rotating, opaque checkerboard was introduced along the line of sight to the pyramid. The checkerboard was introduced at the previously judged distance of the apex of the virtual pyramid. This fact was unknown to all the naive subjects and remained unnoticed throughout

the experiment. Though the virtual pyramid was also presented a second time at the same distance as the first localization, the experimental variations generally concealed this fact from the naive subjects who were led to believe each trial, with or without the checkerboard, involved a potentially different depicted depth. As in Experiment 1, the occasional introduction of unanalyzed sham targets at different depths enforced the naive subjects' belief that the virtual image possibly could be displaced variously in depth for every localization.

3.2 Results

Analysis of the subjects' first localization the virtual object under the three viewing conditions with 2 diopters of accommodative demand closely replicate the findings of Experiment 1. The basic result is a significant 2-way interaction of view condition and age ($F(2,42)=19.160$ $p < 0.0009$) in which age variation effects the judged dis-

tance only for the monocular viewing condition, with the younger subjects (Figure 4).

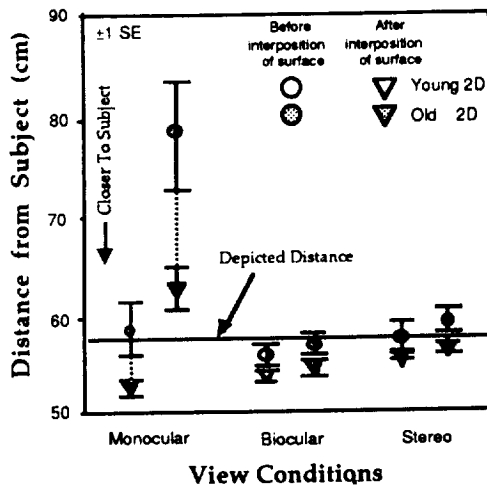


Figure 4: Interaction of Age and Viewing Condition

Analysis of the offset of the mean judged distance to the virtual object associated with the introduction of the physical surface also showed a main effect of viewing condition ($F(2,26) = 91.340$, $p < .0001$) and a significant interaction between viewing condition and age ($F(2,26) = 21.921$, $p < .0001$) (Figure 4). These effects modulated the overall significant offset ($F(1,13) = 90.623$, $p < .0001$) of the judged distance to the virtual object towards the viewer which was caused by introduction of the physical surface. This effect appears for all viewing conditions as a closer localization of the target after interposition of the physical surface.

3.3 Discussion

The first virtual object localization shown in Figure 4 does replicate the 2 D viewing conditions in Experiment 1 showing that the older subjects are unable to use the accommodative information to estimate the virtual object distance and consequently localize the virtual object erroneously towards the background wall. Interestingly, interposition of the checkerboard at the judged distance of the monocularly viewed virtual object causes a substantial forward movement of the judged virtual object position (Figure 4). This change is what would be expected if insertion of the disparity and other cues to nearness of the checkerboard were to cause relative convergence with respect to the position of eyes before its introduction.

It is important to note that the forward movement of the localization is from the *initial judged* position for each appearance of the isolated virtual object. Since the older subjects tended to initially judge the virtual object to be too far away, the tendency for the checkerboard to bring the judge distance closer was corrective

(See Fig. 4). On the other hand, younger subjects viewing monocularly did not significantly misjudge the virtual object distance, as evident from the error bars. For them introduction of the checkerboard was detrimental, causing the virtual object to be judged too near.

In fact, the checkerboard insertion caused the virtual object to generally appear too near for the other two viewing conditions which otherwise supported correct localization of the isolated virtual object. This was generally true for both age groups and suggests that introduction of the physical surface, the checkerboard, could be causing a small relative convergence under these conditions as well. Though disparity information for correct convergence to the virtual object is available under these two conditions, the virtual object providing this information is not of high visual quality. Its presentation corresponds roughly to 20/100 visual acuity. Under such conditions, convergence based on stereoscopic cues might not be precise and could exhibit a fixation disparity. If this error were an exofixation disparity, i.e. it would tend to the distant side, introduction of the high visual fidelity checkerboard could correct it, causing a relative convergence and associated decrease of the judged distance to the virtual object. If this small corrective convergence were incompletely compensated due to the breakdown of distance constancy, errors of localization could be expected. Thus, the change in localization after introduction of the checkerboard could be due to a change in static convergence.

4. Experiment 3

Experiment 3 explicitly tests for such a change associated with the change in judged distance. Attention is focused on the stereoscopic viewing condition to see if the closer judged distance of the virtual object associated with introduction of the physical surface could be associated with an increase in static convergence. Since the amount of expected change is small, for example a change of 3 cm from 58 cm to 55 cm, a sensitive measure of convergence is needed. Angular changes of monocular position of only about 3 arcmin would be expected if the change in localization were explainable by vergence change alone. Such a measurement is difficult to make without encumbering eye tracking technology that preserves a clear visual field. But it is conveniently just at the display resolution of the display configuration used. Therefore, a technique using nonius lines on the display itself was adopted [12]. A nonius line is a line which is broken into two line segments each of which is visible by only one eye. Such lines have typically been used to measure equivalent oculocentric directions to determine the position of the stereoscopic horopter. When the two segments are moved laterally during a period of constant convergence so that they appear to be collinear, their positions may be used to record a specific convergence position.

4.1 Methods

Since pilot studies had suggested that longer distances enhance the effect of introduction of the checkerboard, the virtual pyramid was presented at 108 cm rather than 58 cm away from the subjects' eyes. Such an enhancement was deemed helpful to increase the detectability of any convergence. This display, otherwise similar to that used in Exp. 2, was operated under normal room illumination with one diopter accommodative relief for the virtual image. Flanking nonius lines (lower right panels of Fig. 3) of the same luminance and line width as the pyramid described in the Task section below were also occasionally presented to detect changes in static convergence.

Five men and one woman with measured stereo resolution of better than 1 arcmin participated in the experiment. Some subjects had vision corrected by contact lenses or glasses and were able to wear their corrections during the experiment. Subjects' ages ranged from 17 - 47 and included laboratory personnel as well as paid subjects recruited by a contractor at Ames. Because of the computer control of the experiment it was possible to conduct this experiment double blind.

The subjects task had three basic parts: 1) Localization of an isolated virtual pyramid and measurement of associated static ocular convergence 2) Relocalization of the virtual pyramid in the presence of a real surface either at or in front of the pyramid's apparent distance. 3) Measurement of changes in static convergence associated with the relocalization. The first part of the subject's task was to mechanically place the LED pointer under the nadir of the slowly rotating, wire-frame pyramid. After aligning the pointer, the subjects were presented with two sets of vertical nonius lines just flanking the pyramid (Figure 3, right panel #3). These lines were then adjusted to appear vertically collinear, i.e. to have equal visual directions on each side of the pyramid. This adjustment was made by moving the lower left and right segments with a joystick control (See Figure 3) and effectively recorded the subjects' static convergence during this part of the experiment. Subsequent brief presentations of the nonius line will accordingly show how static convergence may have changed by revealing vertical misalignments. The second part of the task involved another adjustment of the pointer to the pyramid's depth after the slowly, irregularly rotating checkerboard was introduced along the line of sight to the pyramid. The pyramid was then presented a second time at the same depicted depth in this new configuration. As before, the experimental variations generally concealed this fact from the subjects so that they believed each trial, with or without the checkerboard, involved a potentially different depicted depth. And as before, unanalyzed trials with random variation in distance were introduced to maintain the subjects' uncertainty.

After the second judgment of the pyramid's depth, the nonius lines were flashed briefly (ca 250 msec) next to the pyramid while the subjects fixated it (Figure 3, right

panel #4). Then the subjects made a forced choice indicating whether the upper or lower pair of the flashed nonius lines appeared closer laterally. The eye assignments of each segment of the nonius lines were randomly selected so that the meaning of the alternative possibilities in terms of convergence or divergence varied randomly across the trials. The assignment of the lower part of the left nonius line and the upper part of the right line to one eye and the other upper-lower pair to the other eye, produced a differential effect doubling the relative misalignment for any given vergence change and increasing the sensitivity of the technique for detecting changes in convergence. The subject reported by a button press which of the paired nonius lines were closer.

In fact, three different experimental conditions were used in the second part of the experiment because of the need for a control case. In the "on" condition the checkerboard was mechanically introduced at the judged depth of the virtual pyramid object so that the pyramid appeared "on" the checkerboard. For the "in front" condition the checkerboard was introduced 30 cm in front of the judged depth. In the control condition the second judgment was a replication of the first judgment in that the subject made a second judgment of the depth of the virtual object. But this time the subject made the forced-choice judgment of the nonius lines alignment without the addition of the checkerboard. Thus, the control was identical to the experimental conditions except the checkerboard was not introduced into the line of sight. Therefore, this control provides an individual baseline for subjects' judgment biases and changes of their convergence during the course of a trial. Each condition was repeated 15 times for each subject in a randomized block design in which blocks of 5 replications of each condition were repeated. The 6 possible orders of the 3 conditions were distributed randomly across the 6 subjects in the experiment.

The change in judged distance of the virtual object was analyzed in a single factor repeated measures ANOVA. Chi-square analyzes were conducted on each individual subject's distribution of judgments of convergence/divergence for each of the 3 experimental conditions. Taking the control condition as a baseline, the relative strength of convergence could be measured by a ratio of the probability of convergence in each experimental condition to the probability of convergence in each subject's individual control. This ratio allows control for the possibility that subjects might have an individual bias to converge or diverge simply because of a repeated presentation of the virtual object.

4.2 Results

Single factor repeated measures analysis of the effect of superposition of the checkerboard and virtual images replicated the previous observations that the virtual object was moved closer to the viewer ($F(2,10) = 7.549$ $p < 0.01$). Individual data are shown in Figure 5. This effect

was somewhat stronger for the "on" condition than for the "in front" case and varied in strength across the 6 subjects. One subject interestingly showed no major effect.

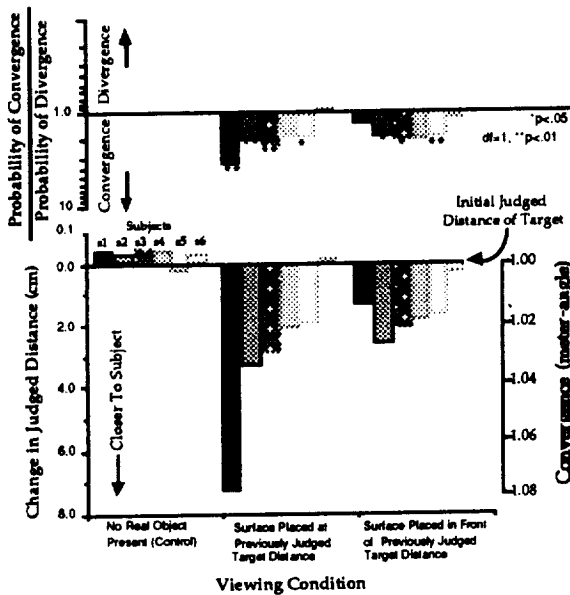


Figure 5: Relative convergence vs. judged distance

The cause of this individual subject's result is illuminated by considering all subjects' tendency to relatively converge during judgment of the depth of the virtual object in the presence of the checkerboard. This tendency is summarized for the experiment in Table 1 which displays the frequency of convergence or divergence indicated by the nonius judgments for all subjects in the three experimental conditions (Chi-square = 20.37, $df=2$, $p<.001$). The control case shows the expected 50:50 break, collapsing across all subjects, while the other two conditions show clear convergence, the "on" condition being somewhat stronger.

	Convergence	Divergence	Total Judgments	Ratio of Convergence to Divergence
On	84	21	105	1.58
In Front	70	35	105	1.32
Control	53	52	105	1.00

Table1. Frequency of convergence to divergence

For further analysis each subject's individual tendency to converge was computed separately as the ratio of their probability of convergence in an exp. condition to their probability of convergence in the control. These ratios are plotted in Fig. 6 for each subject. Since the control was used as reference, all ratios for the control condition are 1. A 2x2 Chi square contingency was also computed to compare the distribution of convergence and divergence for each exp. condition to that of the control. This was done separately for each subject for whom statistically significant differences are indicated by asterisks (Fig. 5).

4.3 Discussion

The individual subject's localization errors in Figure 5 are sorted by the size of the change in the judged position of the virtual object for the "on" condition. These results can then be compared with the ratio of the convergence probabilities. As is clear from the figure, the two measurements are almost perfectly correlated across the subjects. The only subject not to show a displacement of the virtual object caused by the checkerboard, also is the only one to show essentially no relative convergence. The subject showing the largest displacement due to introduction of the checkerboard is also the one with the strongest tendency to converge. The results for the "in front" condition show a weaker apparent displacement of the virtual image but also show a correlation of convergence tendencies and changes in localization. The correlation of relative convergence with magnitude of displacement for the "on" and "in front" conditions across subjects and conditions is; in fact, $r=0.894$ ($df=10$; $t=6.31$, $p<.002$). These results generally support the supposition that the change in judged depth could be related to a change in convergence, but the mechanism underlying this change remains to be clarified. Correction of a fixation disparity, for example, by introduction of a high resolution physical stimulus, for example, is not the only possible mechanism.

One other possibility is that change in convergence is due to so-called perspective [13] or proximal vergence [14]. These phenomena are changes in convergence due to changes in the apparent nearness of objects. They provide evidence that spatial interpretations of the distance of a visual image themselves can simulate the vergence system. Accordingly, the results from the present experiment, while showing that there is a clear oculomotor response associated with the error in judged depth, does not resolve its cause. A change in the apparent nearness of the virtual object due to its appearing to occlude the nearby checkerboard could be the cause of the measured convergence rather than its consequence. This question can only be resolved experimentally.

5. Experiment 4

One approach to analyzing whether the oculomotor effect, i.e. the convergence, observed in Experiment 3 is caused by the superposition of the virtual object on the background is to devise a stimulus condition which on one hand strengthens the oculomotor cues to convergence but on the other hand weakens the visual evidence for occlusion, thus reducing the likelihood of convergence caused by proximal vergence.

We have attempted to create such a stimulus by cutting an annular slot 8 cm wide out of our rotating checkerboard so that the virtual pyramid would be just able to "fall through" the resulting hole (see Figure 6). The

outer rim of the checkerboard was supported by thin radial wire matching the color of the background wall and therefore being invisible to the observers. This stimulus triples the number of moving edges that provide the strong disparity discontinuity which could be the stimulus to convergence that could be the cause of the change in static convergence observed in Experiment 3. If the better stimulus to convergence provided by the checkerboard were the cause of the change in static convergence, this stimulus should strengthen the effect. On the other hand, the slotted hole virtually eliminates the visual evidence for occlusion. If proximal vergence triggered by occlusion were the cause of the change in judged distance, one would expect not to find a change in the judged distance of the virtual object when it is presented in the slot.



Figure 6. Slotted physical surface

5.1 Methods

Nine subjects aged 23 - 47 who were either laboratory personnel or paid subjects provided by Bionetics were used in Experiment 4.

This experiment was conducted using a methodology equivalent to that of Experiment 2 for stereoscopic virtual objects. Two different depicted distances of the virtual objects, 83 cm and 108 cm, were randomly ordered into blocks of 20 runs. In each block either a solid checkerboard or a slotted checkerboard was introduced along the line of sight of the virtual object. After introduction of the checkerboard, the change in the judged distance to the virtual object was measured. Block types were alternated for all subjects. All but one subject, from whom half of his data was lost, experienced 4 blocks, making a total of 80 judgments per subject. The order of presentation of the two checkerboard types was counterbalanced across subjects.

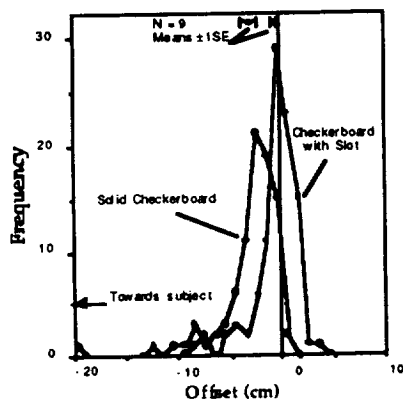


Figure 7. Hole effect on judged distance to virtual objects.

5.2 Results

Analysis of variance showed that while the solid disk caused a previously observed offset of the judged virtual object distance towards the observer (Mean: 2.80 cm; SE: ± 0.75 cm), the slotted disk caused only a mean 0.72 cm (SE ± 0.49 cm) change. This difference was statistically significant. $F(1,8) = 19.605$ $p < 0.002$. The offset for the slotted disk condition, though quite close to 0, is statistically significantly less than 0. ($t = -2.59$, $df = 8$, $p < 0.04$)

The size of the offset was larger for the greater depicted distance. The offset for the 83 cm virtual object was 1.49 cm (SE ± 0.43) and 2.03 cm (SE ± 0.57) for the 108 cm object. This difference was just significant ($F(1,8) = 5.349$ $p < 0.05$). There was no statistical interaction ($F(1,8) = 3.068$ $p > 0.05$) so the effects of placing the virtual object in the slot was statistically indistinguishable for the two presentation distances. Accordingly, the data from the two presentation distances may be collapsed as in Figure 7. This figure shows the full distribution of all the subjects responses illustrating the effect of the slot on the judged distance to the virtual object.

5.3 Discussion

As is clear from Figure 7, introduction of the slot in the checkerboard that removed the occlusion between the virtual object and the checkerboard greatly reduced the offset in judged virtual object position produced by the checkerboard introduction. This reduction makes the oculomotor explanation of the change less likely since the binocular depth cues of the checkerboard would be expected, if anything, to strengthen any vergence response. Since the overlap of the virtual object contours and those of the checkerboard seems to be the key feature causing the shift in its judged position, the occlusion cue placing the virtual object in front of the checkerboard seems to be the best explanation for the nearer localization of the virtual objects. The change in static convergence, thus, appears to be a consequence of proximal vergence and resembles effects in recent reports of convergence being driven by the kinetic depth effects which produce a perception of nearness [15].

There is, however, another possible interpretation of the results which could be based on the proposal that the alternative explanations of the offset of the judged distance of the virtual object are not mutually exclusive [16]. In this view the introduction of the slotted checkerboard stimulus could introduce a proximal vergence tendency to fixate farther away since any objects seen through the slot would necessarily be more distant than the slot. Were this tendency for a more distant fixation to occur, it would oppose any tendency of the visual information to reduce oculomotor bias. The net effect of the opposition could account for the almost negligible forward displacement of the virtual object observed while subjects were exposed to the slotted checkerboard.

One way to dissociate the two influences could be to study the individual oculomotor biases of each subject, measuring for example their phoria with a Maddox Rod test or equivalent. To the extent subjects exhibit exophoria, one could expect a exofixation disparity while viewing the virtual object. Its removal by presentation of the checkerboards could be associated with the change in judged distance. An explanation of the observed effects of the checkerboards solely in terms of proximal vergence would not predict a correlation between the phoria tests and the subjects' individual phoria variations. Accordingly, future investigations could examine individual subject's oculomotor biases as a technique to more precisely determine the cause of the offset of the change in judged distance associated with superimposition of virtual objects against physical surfaces.

6. Design Considerations

1) The present results were observed with a static eye point and can be expected to change when significant lateral head-movement producing motion parallax is introduced. Nevertheless, it is important to realize that since many of the new applications of head-mounted see-through displays, in fact, will involve relatively static viewing, the conditions used remain practically relevant.

2) Since weight and cost considerations may argue for the use of monocular displays, they are likely to be initial candidates for many applications. Accordingly, such displays should have a variable focus control to appropriately direct the convergence of prepresbyopic users. Designers and supervisors should be aware that operators over 40 will generally not benefit from the variable focus adjustment.

3) Biocular and stereo displays should be used with a bore-sighting procedure in which focus is adjusted to a reference target so as to correct for any errors in depth due to inappropriate vergence.

4) Computer generated or other targets presented binocularly should have individually tailored stereo disparity to correct their spatial localization so as to compensate for the tendency of virtual objects to appear to float in front of the surfaces that they are seen against.

7. Acknowledgments

Earlier reports of the above: 1995, 1996 Proc. HFES, 1995 Psychonomics Soc., the 1995 IFAC Proc. A more detailed report will appear in *Human Factors*.

8. References

[1] Azuma, R. & Bishop, G. (1994) Improving static and dynamic registration in an optical see-through HMD. Proc. of SIGGRAPH '94, (pp. 197-204) July 24-29, Orlando, FL, New York: ACM.

[2] Janin, A.L., Mizell, D.W., & Caudell, T.P. (1993) Calibration of head-mounted displays for augmented reality applications. Proc. of IEEE VRAIS '93 (pp. 246-255), Seattle New York: IEEE.

[3] Rolland, J.P., Ariely, D., & Gibbon, W. (1995) Towards quantifying depth and size perception in 3D virtual environments. *Presence* 4 (1), 24-49.

[4] Ellis, Stephen R. & Menges, Brian M. (1997) Judged distance to virtual objects in the near visual field *Presence* 6(4), 452-460.

[5] Moses, Robert A. (1987) Accommodation. In R. A. Moses, & W. M. Haret Jr. *Adler's physiology of the eye*, (pp. 291-310). Washington D. C. : Mosby.

[6] Gogel, W.C. & Tietz, J.D. (1973) Absolute motion parallax and the specific distance tendency. *Perception and Psychophysics*, 13: 284-292.

[7] Owens, D. A. & Liebowitz, H. W. (1976) The specific distance tendency. *Perception and Psychophysics*, 20, 2-9.

[8] Owens, D. A. & Liebowitz, H. W. (1983) Perceptual and motor consequences of tonic vergence. In Eds. K. J. Ciuffreda & C. M. Shor *Vergence eye movements: basic and clinic aspects* (p. 50), Boston: Butterworths.

[9] Zuber, B. (1965) Physiological control of eye movements in humans. (p.103). Ph.D. thesis, Cambridge, Mass: MIT.

[10] Brown, J. L. (1965) Afterimages In C. H. Graham, (Ed.) *Vision & Visual Perception*, (p. 485). New York: Wiley

[11] Ellis, Stephen R. & Menges, Brian M. (1997) Effects of age on the judged distance to virtual objects in the near visual field, In Wendy. A. Rogers, (Ed.) *Designing for an aging population*: (pp. 15-19). Santa Monica, CA: Human Factors and Ergonomics Society.

[12] Ellis, S. R. Bucher, U. J. & Menges, B. M. (1995) The relationship of binocular convergence to error in the judged distance of virtual objects. Proceedings of the International Federation for Automatic Control, (pp. 297-301) Boston, June 26-27.

[13] Enright, J.T. (1991) Paradoxical monocular stereopsis and perspective vergence. In S. R. Ellis, et al. (Eds). *Pictorial communication in virtual and real environments*. (pp. 567-576). London: Taylor and Francis.

[14] Ciuffreda, K. J. (1992) Components of clinical near vergence testing. *Journal of Behavioral Optometry*, 3, 1, 3-13.

[15] Ringach, D. L., Hawken, M. J. & Shapley, R. (1997) Binocular eye movements caused by the perception of three dimensional structure from motion. *Vision Research*, 36, (10) 1479-1492.

[16] Shebilske, W. (1997) personal communication.

